Precoded Integer-Forcing Equalization Universally Achieves the MIMO Capacity up to a Constant Gap

Or Ordentlich and Uri Erez

***Abstract*—A single-user multiple-input multiple-output communication scheme is considered where a transmitter, equipped with multiple antennas, encodes the data into independent streams via a single linear code. The coded streams are then precoded using linear space-time modulation from the class of perfect codes. At the receiver side, integer-forcing equalization is applied, followed by standard single-stream decoding. It is shown that this communication architecture achieves the capacity of any multiple-input multiple-output channel up to a constant gap.**

I. Introduction

The Gaussian Multiple-Input Multiple-Output (MIMO) channel has been the focus of extensive research efforts since the pioneering works of Foschini [1], Foschini and Gans [2], and Telatar [3]. While the capacity of the channel, under various assumptions on channel state information, is easy to derive, practical schemes that allow to approach the capacity are only known for the two extremes: when the channel is either ergodic, with or without channel state information at the transmitter, or static where the transmitter knows the channel matrix. In contrast, this paper considers a static (non-ergodic) scenario where the receiver has full channel knowledge whereas the transmitter either has no knowledge of the channel, or alternatively knows only its capacity.

More specifically, a single-user complex MIMO channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} \tag{1}$$

with M transmit and N receive antennas is considered. The input vector $\mathbf{x}$ is subject to the power constraint[1]

$$\mathbb{E}(\mathbf{x}^\dagger \mathbf{x}) \leq M \cdot \mathsf{SNR},$$

and the additive noise $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is a circularly symmetric complex normal vector with zero mean and unit variance.

The mutual information of this channel is maximized using a circularly symmetric complex Gaussian input with zero mean and covariance matrix $\mathbf{Q}$ satisfying $\mathrm{trace}(\mathbf{Q}) \leq M \cdot \mathsf{SNR}$, and is given by [3]

$$C = \max_{\mathbf{Q} \succ 0 \,:\, \mathrm{trace}\,\mathbf{Q} \leq M \cdot \mathsf{SNR}} \log\det\left(\mathbf{I} + \mathbf{Q}\mathbf{H}^\dagger \mathbf{H}\right). \tag{2}$$

The choice of $\mathbf{Q}$ that maximizes (2) is determined by the water-filling solution. When the matrix $\mathbf{H}$ is known at both transmission ends, i.e., in a closed-loop scenario, this mutual information is the capacity of the channel and may closely be approached using the singular-value decomposition in conjunction with standard scalar codes designed for an additive white Gaussian noise (AWGN) channel. Often, the sub-optimal choice $\mathbf{Q} = \mathsf{SNR} \cdot \mathbf{I}$ is used, resulting in the *white-input (WI) mutual information*

$$C_{\text{WI}} = \log\det\left(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger \mathbf{H}\right).$$

For all channel matrices $\mathbf{H}$ and all values of SNR, the WI mutual information loses less than $M \log M$ bits w.r.t. capacity [4], [5]. Thus, a scheme that performs a constant gap from C_{WI} also performs a constant gap from the closed-loop capacity C. In this paper, an open-loop setting is considered. In this case, transmitting a white input is a natural choice and C_{WI} will serve as a benchmark in the sequel.

While the theoretical performance limits of open-loop communication over a Gaussian MIMO channel are well understood, unlike for closed-loop transmission, much is still lacking when it comes to practical schemes that are able to approach these limits. Such schemes are known for the 1×2 MISO channel where Alamouti modulation offers an optimal solution. More generally, modulation via orthogonal space-time block "codes" allows to approach the WI mutual information using scalar AWGN coding and decoding in the limit of small rate, where the mutual information is governed solely through the Frobenius norm of the channel matrix.

Beyond the low rate regime, i.e., when the multiple degrees of freedom offered by the channel need to be utilized, despite considerable work and progress, much remains to be hoped for. Since finding a practical scheme that is able to approach the WI mutual information with no transmitter channel knowledge turned out to be a very challenging task, less demanding benchmarks became accepted in the literature. First, since statistical modeling of a wireless communication link is often available, one may be content with guaranteeing good performance only for channel realizations that have a "high" probability. Moreover, to further simplify analysis and design, the asymptotic criterion of the diversity-multiplexing tradeoff [4] has widely been adopted. Unfortunately, such characterizations offer only a coarse figure of merit for assessing schemes.

Realizing that a metric based on a particular statistical model is prone to be non-robust to these very assumption, a major achievement was the definition of approximately universal space-time codes by Tavildar and Vishwanath [6] and in identifying and developing new schemes that satisfy this property [7]–[9]. Roughly speaking, approximate-universality

The work of U. Erez was supported in part by the Israel Science Foundation under Grant No. 1557/13. The work of O. Ordentlich was supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, a fellowship from The Yitzhak and Chaya Weinstein Research Institute for Signal Processing at Tel Aviv University and the Feder Family Award.

O. Ordentlich and U. Erez are with Tel Aviv University, Tel Aviv, Israel (email: ordent,uri@eng.tau.ac.il).

[1]In this paper $(\mathbf{x})^\dagger$ is the conjugate transpose of $\mathbf{x}$.

guarantees that a scheme is DMT optimal for any statistical channel model.

Approximately universal schemes still suffer from the asymptotic nature of the DMT criterion. In particular, since uncoded transmission suffices to achieve the optimal DMT, knowing that a scheme is DMT optimal does not provide any performance guarantee for a channel with fixed (and even known) capacity.

In this work, a practical communication architecture that achieves the capacity of any MIMO channel up to a constant gap is studied. Specifically, in the considered scheme, the transmitter encodes the data into independent streams via a single linear code. The coded streams are then linearly precoded using the generating matrix of a space-time code from the class of perfect codes [8]–[11]. At the receiver side, integer-forcing equalization [12] is applied.

Integer-forcing equalization essentially reduces to lattice-reduction (LR) in the case of uncoded transmission. Lattice-reduction aided receivers for perfect space-time modulated QAM constellations were considered in the literature, and were shown to be DMT optimal [13]. The key difference is that while the latter approach involves uncoded transmission and symbol-by-symbol detection, the proposed architecture uses linearly coded streams and the detection phase is replaced with equalization. This is turn leads to performance guarantees that are valid at any (fixed) transmission rate.

II. Preliminaries on Integer-Forcing Equalization

Integer-Forcing equalization is a low-complexity architecture for the MIMO channel, which was proposed by Zhan *et al*. [12]. The key idea underlying IF is to first decode linear combinations of the signals transmitted by all antennas, and then, after the noise is removed, invert those linear combinations to recover the individual transmitted signals. This is made possible by transmitting codewords from the *same* linear/lattice code from all M transmit antennas, leveraging the property that linear codes are closed under (modulo) linear combinations with integer-valued coefficients.

A. Nested Lattice Codes

Let $\Lambda_c \subset \Lambda_f$ be a pair of n-dimensional nested lattices (see [14] for a rigorous treatment of lattice definitions and properties). The lattice Λ_c is referred to as the coarse lattice and Λ_f as the fine lattice. Denote by $\mathcal{V}_c$ the Voronoi region of Λ_c, and define the second moment of Λ_C as

$$\sigma^2(\Lambda_c) \triangleq \frac{1}{n}\frac{1}{\mathrm{Vol}(\mathcal{V}_c)}\int_{\mathbf{u}\in\mathcal{V}_c} \|\mathbf{u}\|^2 d\mathbf{u}.$$

A nested lattice codebook $\mathcal{C} = \Lambda_f \cap \mathcal{V}_c$, with rate[2]

$$R = \frac{1}{n}\log|\Lambda_f \cap \mathcal{V}_c| \,\frac{\text{bits}}{\text{channel use}}$$

is constructed from the nested lattice pair. The codebook is scaled such that $\sigma^2(\Lambda_c) = \mathsf{SNR}/2$.

[2]All logarithms in this paper are to the base 2.

Example 1: Three examples of common nested lattice codebooks are given below. More examples can be found in [15].

- *Uncoded transmission* - The simplest nested lattice codebook is an uncoded one, where the fine lattice Λ_f is the integer lattice $\mathbb{Z}$ whereas the coarse lattice is $\Lambda_c = q\mathbb{Z}$ for some integer $q > 1$. The Voronoi region in this case is $\mathcal{V}_c = [-q/2, q/2)$ and the obtained nested lattice codebook $\mathcal{C}$ consists of all integers in the interval $[-q/2, q/2)$. The rate of this codebook is $R = \log q$ bits/channel use.
- *q-ary linear code without shaping* - A more sophisticated, yet reasonable to implement, nested lattice codebook can be obtained by lifting a q-ary linear code with block length n to Euclidean space using Construction A [16], [17], and taking the resulting lattice as Λ_f. The coarse lattice is taken as $\Lambda_c = q\mathbb{Z}^n$, as in the uncoded case. The obtained nested lattice codebook $\mathcal{C}$ is therefore simply the q-ary linear code coupled with a PAM constellation.
- *"Good" nested lattice pair of high dimension* - A third option is to use a pair of lattices of high dimension where the fine lattice is "good" for coding over an AWGN channel, whereas the coarse lattice is "good" for mean squared error quantization (see [14] for precise definitions of "goodness"). The obtained nested lattice codebook admits a relatively simple performance analysis, that yields closed-form rate expressions. However, implementing such a codebook is more complicated (although some progress in this direction was made in [18]).
  The performance improvement obtained by using such a codebook w.r.t. a q-ary linear code without shaping is bounded by $1/2\log(2\pi e/12)$ bits per real dimension, provided that the q-ary linear code performs well over an AWGN channel.

B. Description of IF equalization

In the IF scheme, the information bits to be transmitted are partitioned into $2M$ streams, labeled $\{1_\Re, 1_\Im, \ldots, M_\Re, M_\Im\}$. Each of the $2M$ streams is encoded by the nested lattice code $\mathcal{C}$. In particular, the stream $m_\Re$, consisting of nR information bits, is mapped to a lattice point $\mathbf{t}_{m_\Re} \in \mathcal{C}$. Then, a random dither $\mathbf{d}_{m_\Re} \in \mathbb{R}^{1\times n}$ uniformly distributed over $\mathcal{V}_c$ and statistically independent of $\mathbf{t}_{m_\Re}$, known to both the transmitter and the receiver, is used to produce the signal

$$\mathbf{x}_{m_\Re} = [\mathbf{t}_{m_\Re} - \mathbf{d}_{m_\Re}] \bmod \Lambda_c.$$

The signal $\mathbf{x}_{m_\Re}$ is uniformly distributed over $\mathcal{V}_c$ and is statistically independent of $\mathbf{t}_{m_\Re}$ due to the Crypto Lemma [14, Lemma 1]. It follows that

$$\frac{1}{n}\mathbb{E}\|\mathbf{x}_{m_\Re}\|^2 = \sigma^2(\Lambda_c) = \frac{\mathsf{SNR}}{2}.$$

A similar procedure is used to construct the signal $\mathbf{x}_{m_\Im}$. The mth antenna transmits the signal $\mathbf{x}_m = \mathbf{x}_{m_\Re} + i\mathbf{x}_{m_\Im} \in \mathbb{C}^{1\times n}$ over n consecutive channel uses.

Let $\mathbf{X} = [\mathbf{x}_1^T \;\cdots\; \mathbf{x}_M^T]^T \in \mathbb{C}^{M\times n}$. The received signal is

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z},$$

where $\mathbf{Z} \in \mathbb{C}^{N\times n}$ is a vector with i.i.d. circularly symmetric complex normal entries. Letting the subscripts $\Re$ and $\Im$ denote the real and imaginary parts of a matrix, respectively, the channel can be expressed by its real-valued representation

$$\begin{bmatrix} \mathbf{Y}_\Re \\ \mathbf{Y}_\Im \end{bmatrix} = \begin{bmatrix} \mathbf{H}_\Re & -\mathbf{H}_\Im \\ \mathbf{H}_\Im & \mathbf{H}_\Re \end{bmatrix} \begin{bmatrix} \mathbf{X}_\Re \\ \mathbf{X}_\Im \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_\Re \\ \mathbf{Z}_\Im \end{bmatrix},$$

which will be written as $\tilde{\mathbf{Y}} = \tilde{\mathbf{H}}\tilde{\mathbf{X}} + \tilde{\mathbf{Z}}$ for notational compactness. Let $\tilde{\mathbf{T}} = [\mathbf{t}_{1_\Re}^T \ \cdots \ \mathbf{t}_{M_\Re}^T \ \mathbf{t}_{1_\Im}^T \ \cdots \ \mathbf{t}_{M_\Im}^T]^T$ be a $2M \times n$ matrix whose rows consist of the lattice points representing each of the $2M$ bit streams, and $\tilde{\mathbf{D}} = [\mathbf{d}_{1_\Re}^T \ \cdots \ \mathbf{d}_{M_\Re}^T \ \mathbf{d}_{1_\Im}^T \ \cdots \ \mathbf{d}_{M_\Im}^T]^T$ be a $2M \times n$ matrix whose rows correspond to the $2M$ different dither vectors.

The IF receiver chooses an equalizing matrix $\mathbf{B} \in \mathbb{R}^{2M\times 2N}$ and a full-rank target integer-valued matrix $\mathbf{A} \in \mathbb{Z}^{2M\times 2M}$, and computes

$$\begin{aligned} \tilde{\mathbf{Y}}_{\text{eff}} &= \left[\mathbf{B}\tilde{\mathbf{Y}} + \mathbf{A}\tilde{\mathbf{D}}\right] \bmod \Lambda_c \\ &= \left[\mathbf{A}\tilde{\mathbf{T}} + (\mathbf{B}\tilde{\mathbf{H}} - \mathbf{A})\tilde{\mathbf{X}} + \mathbf{B}\tilde{\mathbf{Z}}\right] \bmod \Lambda_c \\ &= [\mathbf{V} + \mathbf{Z}_{\text{eff}}] \bmod \Lambda_c, \end{aligned} \tag{3}$$

where $\mathbf{V} \triangleq \left[\mathbf{A}\tilde{\mathbf{T}}\right] \bmod \Lambda_c$ is a $2M \times n$ matrix with each row being a codeword in $\mathcal{C}$ owing to the linearity of the code, $\mathbf{Z}_{\text{eff}} \triangleq (\mathbf{B}\tilde{\mathbf{H}} - \mathbf{A})\tilde{\mathbf{X}} + \mathbf{B}\tilde{\mathbf{Z}}$ is additive noise statistically independent of $\mathbf{V}$ (as $\tilde{\mathbf{X}}$, as well as $\tilde{\mathbf{Z}}$ are statistically independent of $\tilde{\mathbf{T}}$), and the notation $\bmod \Lambda_c$ is to be understood as reducing *each row* of the obtained matrix modulo the coarse lattice. Thus, each row of $\tilde{\mathbf{Y}}_{\text{eff}}$ is the output of a modulo additive noise channel whose input is a codeword from $\mathcal{C}$

$$\tilde{\mathbf{y}}_{\text{eff},k} = [\mathbf{v}_k + \mathbf{z}_{\text{eff},k}] \bmod \Lambda_c, \quad k = 1, \ldots, 2M. \tag{4}$$

It is evident that the IF receiver transforms the original MIMO channel into a set of $2M$ parallel[3] point-to-point sub-channels. The output of each sub-channel can be decoded separately. If the decoding is successful over all $2M$ sub-channels, the receiver has access to $\mathbf{V}$, from which it can recover the matrix $\tilde{\mathbf{T}}$ by solving the (modulo) set[4] of equations $\mathbf{V} \triangleq \left[\mathbf{A}\tilde{\mathbf{T}}\right] \bmod \Lambda_c$. See Figures 1 and 2.

Define the effective variance of $\mathbf{z}_{\text{eff},k}$ as

$$\begin{aligned} \sigma^2_{\text{eff},k} &\triangleq \frac{1}{n}\mathbb{E}\left\|\mathbf{z}_{\text{eff},k}\right\|^2 \\ &= \frac{1}{n}\mathbb{E}\left\|(\mathbf{b}_k\tilde{\mathbf{H}} - \mathbf{a}_k)\tilde{\mathbf{X}} + \mathbf{b}_k\tilde{\mathbf{Z}}\right\|^2 \\ &= \frac{\mathsf{SNR}}{2}\|(\mathbf{b}_k\tilde{\mathbf{H}} - \mathbf{a}_k)\|^2 + \frac{1}{2}\|\mathbf{b}_k\|^2, \end{aligned}$$

where $\mathbf{b}_k$ and $\mathbf{a}_k$ are the kth rows of $\mathbf{B}$ and $\mathbf{A}$, respectively. A natural criterion for choosing the equalizing matrix $\mathbf{B}$ and the target integer-valued matrix $\mathbf{A}$ is to minimize the effective noise variances. It turns out [12] that for a given matrix $\mathbf{A}$, the optimal choice of $\mathbf{B}$ under this criterion is

$$\mathbf{B}^{\text{opt}} = \mathbf{A}\tilde{\mathbf{H}}^T\left(\frac{1}{\mathsf{SNR}}\mathbf{I} + \tilde{\mathbf{H}}\tilde{\mathbf{H}}^T\right)^{-1},$$

which results in the effective variances

$$\sigma^2_{\text{eff},k} = \frac{\mathsf{SNR}}{2}\mathbf{a}_k^T\left(\mathbf{I} + \mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\mathbf{a}_k, \tag{5}$$

for $k = 1, \ldots, 2M$.

Define the effective signal-to-noise ratio (SNR) at the kth sub-channel as

$$\begin{aligned} \mathsf{SNR}_{\text{eff},k} &\triangleq \frac{\sigma^2(\Lambda_c)}{\sigma^2_{\text{eff},k}} \\ &= \frac{\frac{\mathsf{SNR}}{2}}{\frac{\mathsf{SNR}}{2}\mathbf{a}_k^T\left(\mathbf{I} + \mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\mathbf{a}_k} \\ &= \left(\mathbf{a}_k^T\left(\mathbf{I} + \mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\mathbf{a}_k\right)^{-1}, \end{aligned}$$

and let

$$\mathsf{SNR}_{\text{eff}} \triangleq \min_{k=1,\ldots,2M} \mathsf{SNR}_{\text{eff},k}. \tag{6}$$

For IF equalization to be successful, decoding over all $2M$ sub-channels should be correct. Therefore, the worst sub-channel constitutes a bottleneck. For this reason, the total performance of the equalizer is dictated by $\mathsf{SNR}_{\text{eff}}$.

C. Performance of the IF Equalizer

When the codebook $\mathcal{C}$ is constructed from a good pair of nested lattices (see Example 1) the distribution of the effective noise at each sub-channel k, which is a linear combination of an AWGN and $2M$ dither vectors, approaches that of an AWGN with zero mean and variance $\sigma^2_{\text{eff},k}$ [21]. Good nested lattice codebooks can achieve any rate satisfying

$$R < \frac{1}{2}\log\left(\mathsf{SNR}_{\text{eff},k}\right) \tag{7}$$

over an AWGN modulo channel with signal-to-noise ratio $\mathsf{SNR}_{\text{eff},k}$ [14], [21]. Since $\mathbf{v}_k$ is a codeword from a good nested lattice code and $\mathbf{z}_{\text{eff},k}$ approaches an AWGN in distribution, $\mathbf{v}_k$ can be decoded [12], [21] from $\tilde{\mathbf{y}}_{\text{eff},k}$ as long as the rate of the codebook $\mathcal{C}$ satisfies (7). It follows that as long as

$$R < \frac{1}{2}\log\left(\mathsf{SNR}_{\text{eff}}\right),$$

all sub-channels $k = 1, \ldots, 2M$ can decode their linear combinations $\mathbf{v}_k$ without error, and therefore IF equalization can achieve any rate satisfying

$$\begin{aligned} R_{\text{IF}} &< 2M\frac{1}{2}\log\left(\mathsf{SNR}_{\text{eff}}\right) \\ &= M\log\left(\mathsf{SNR}_{\text{eff}}\right). \end{aligned} \tag{8}$$

As mentioned in Example 1, good nested lattice codebooks may be difficult to implement in practice. A more appealing alternative is therefore to use a q-ary linear code without shaping. In this case, the effective noise $\mathbf{z}_{\text{eff},k}$ at each sub-channel is a linear combination of an AWGN and $2M$ random

[3] The additive noise vectors $\mathbf{z}_{\text{eff},1}, \ldots, \mathbf{z}_{\text{eff},2M}$ are not statistically independent. Thus, strictly speaking, the $2M$ effective channels $\tilde{\mathbf{y}}_{\text{eff},1}, \ldots, \tilde{\mathbf{y}}_{\text{eff},2M}$ are not parallel. However, the IF decoder ignores the correlation between the noise vectors and treats the $2M$ obtained channels as being parallel. Some improvement can be obtained by exploiting these correlations [19]. Yet, we disregard this possibility in the present paper.

[4] In [20] it is shown that it suffices that $\mathbf{A}$ is invertible over $\mathbb{R}$ in order to recover $\tilde{\mathbf{T}}$ from $\mathbf{V}$.

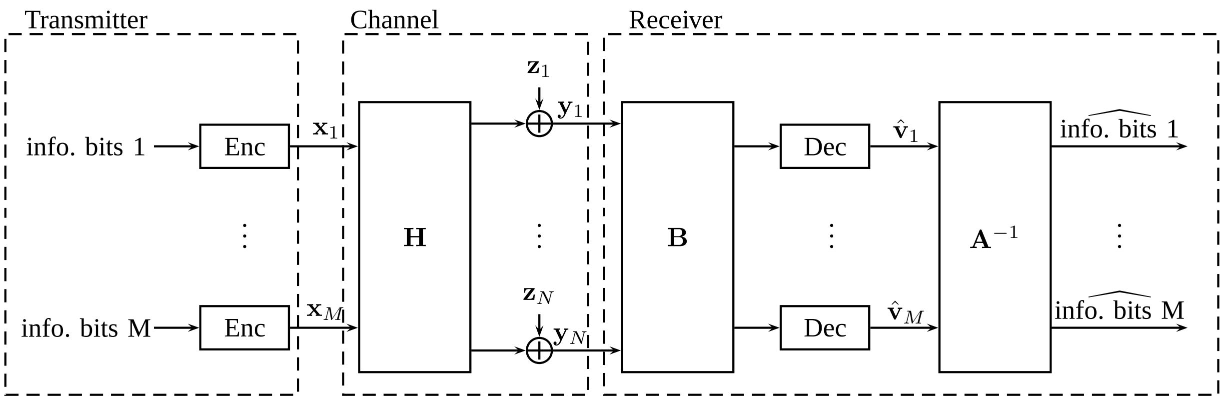

Fig. 1. A schematic overview of the integer-forcing transmitter and receiver. For simplicity, the dithers are not depicted in the figure, and a real-valued channel is assumed. At the transmitter, the information bits are split to M streams. Each stream is encoded by the same similar linear codebook and transmitted by one of the transmit antennas. The receiver first applies the equalizing matrix $\mathbf{B}$ which is supposed to equalize the channel $\mathbf{H}$ to an equivalent channel with transfer matrix approximately equal to $\mathbf{A}$. The equalizer produces M outputs, each of which is an integer-valued linear combination of the transmitted codewords plus effective noise. Each one of these outputs is decoded separately, and finally the outputs of the M decoders are multiplied by $\mathbf{A}^{-1}$ to produce the transmitted codewords. The codewords are then mapped to information bits (this step is not depicted in the figure).

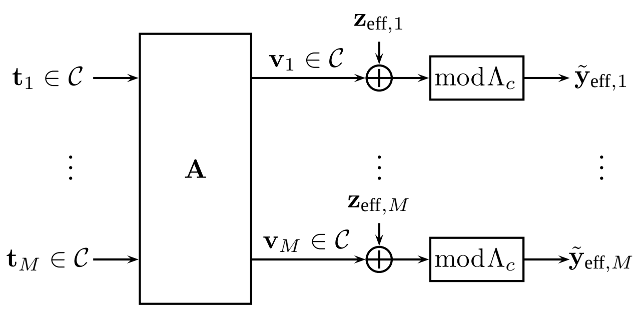

Fig. 2. An illustration of the effective channel obtained when integer-forcing equalization is used. The effective channel consists of M parallel sub-channels. The output of each sub-channel is an integer-valued linear combination of lattice points, which is itself a lattice point, plus effective noise, modulo the coarse lattice Λ_c.

dithers uniformly distributed over the Voronoi region of a 1-D integer lattice. This effective noise is i.i.d. (in contrast to the case where a higher-dimensional coarse lattice is used where $\mathbf{z}_{\text{eff},k}$ has memory). It was shown in [22] that, for a prime q large enough, q-ary linear codes without shaping can achieve any rate satisfying

$$R < \frac{1}{2}\log\left(\mathsf{SNR}_{\text{eff}}\right) - \frac{1}{2}\log\left(\frac{2\pi e}{12}\right)$$

over a modulo channel with additive effective noise $\mathbf{z}_{\text{eff},k}$. Therefore, IF equalization using q-ary linear codes without shaping can achieve any rate below

$$R_{\text{IF,q-ary}} < M\log\left(\mathsf{SNR}_{\text{eff}}\right) - M\log\left(\frac{2\pi e}{12}\right) \tag{9}$$

When a specific q-ary linear code (such as an LDPC code or a turbo code) is used, the achievable rate is further degraded by $2M$ times the code's gap-to-capacity at the target error probability.

Finally, consider the case of uncoded transmission. In this case, $\Lambda_f = \gamma\mathbb{Z}$ and $\Lambda_c = \gamma q\mathbb{Z}$, where $\gamma = \sqrt{12\mathsf{SNR}/(2q^2)}$ is chosen such as to meet the power constraint, and $q > 1$ is an integer (see Example 1). The output of the kth sub-channel with uncoded transmission is

$$\begin{aligned}\tilde{y}_k &= [v_k + z_{\text{eff},k}] \mod \gamma q\mathbb{Z}\\ &= [\gamma u_k + z_{\text{eff},k}] \mod \gamma q\mathbb{Z},\\ &= \left[\gamma\left(u_k + \frac{1}{\gamma}z_{\text{eff},k}\right)\right] \mod \gamma q\mathbb{Z},\\ &= \left[u_k + \frac{1}{\gamma}z_{\text{eff},k}\right] \mod q\mathbb{Z}\end{aligned}$$

where $u_k = v_k/\gamma \in \mathbb{Z}$. Thus, a detection error occurs only if

$$\left|\frac{1}{\gamma}z_{\text{eff},k}\right| \geq \frac{1}{2}.$$

In order to bound the probability of detection error at the kth sub-channel, a simple lemma, which is based on [15, Theorem 7] is needed.

Lemma 1: Consider the random variable

$$z_{\text{eff}} = \sum_{\ell=1}^{L}\alpha_\ell z_\ell + \sum_{k=1}^{K}\beta_k d_k$$

where $\{z_\ell\}_{\ell=1}^L$ are i.i.d. Gaussian random variables with zero mean and some variance σ_z^2 and $\{d_k\}_{k=1}^K$ are i.i.d. random variables uniformly distributed over the interval $[-\rho/2, \rho/2)$ for some $\rho > 0$. The random variables $\{z_\ell\}_{\ell=1}^L$ are statistically independent of $\{d_k\}_{k=1}^K$. Let $\sigma_{\text{eff}}^2 \triangleq \mathbb{E}(z_{\text{eff}}^2)$. Then

$$\Pr(z_{\text{eff}} > \tau) = \Pr(z_{\text{eff}} < -\tau) \leq \exp\left\{-\frac{\tau^2}{2\sigma_{\text{eff}}^2}\right\}.$$

Proof: See Appendix A. ■

Now, using Lemma 1, the probability of detection error at

the kth sub-channel can be bounded as

$$\begin{aligned} P_{e,k} &\triangleq \Pr\left(\hat{v}_k \neq v_k\right) \\ &\leq \Pr\left(\left|\frac{1}{\gamma} z_{\text{eff},k}\right| \geq \frac{1}{2}\right) \\ &\leq 2\exp\left\{-\frac{\gamma^2}{8\sigma^2_{\text{eff},k}}\right\} \\ &\leq 2\exp\left\{-\frac{3\mathsf{SNR}}{4q^2\sigma^2_{\text{eff},k}}\right\} \\ &\leq 2\exp\left\{-\frac{3}{2}\frac{1}{q^2}\mathsf{SNR}_{\text{eff},k}\right\}, \end{aligned}$$

where the definition of $\mathsf{SNR}_{\text{eff},k}$ was used in the last inequality. Using the fact that $q = 2^R$ and that $\mathsf{SNR}_{\text{eff},k} \geq \mathsf{SNR}_{\text{eff}}$ for all $k = 1, \ldots, 2M$, the detection error probability at each of the $2M$ sub-channels can be further bounded as

$$P_e \leq 2\exp\left\{-\frac{3}{2}2^{2\left(\frac{1}{2}\log(\mathsf{SNR}_{\text{eff}})-R\right)}\right\}.$$

Since the IF equalizer makes an error only if a detection error occurred in at least one of the $2M$ sub-channels, and since the total transmission rate is $R_{\text{IF}} = 2MR$, the total error probability of the IF equalizer with uncoded transmission rate R_{IF} is bounded by

$$\begin{aligned} P_{e,\text{IF-uncoded}} &\leq 4M\exp\left\{-\frac{3}{2}2^{2\left(\frac{1}{2}\log(\mathsf{SNR}_{\text{eff}})-\frac{R_{\text{IF}}}{2M}\right)}\right\} \\ &= 4M\exp\left\{-\frac{3}{2}2^{\frac{1}{M}\left(M\log(\mathsf{SNR}_{\text{eff}})-R_{\text{IF}}\right)}\right\}. \end{aligned} \tag{10}$$

Remark 1: Integer forcing equalization with uncoded transmission is quite similar to the extensively studied lattice-reduction-aided linear decoders framework [13], [23], [24]. However, two subtle differences should be pointed out. First, under the framework of LR-aided linear decoding, the target integer matrix $\mathbf{A}$ has to be unimodular, i.e., it has to satisfy $|\det(\mathbf{A})| = 1$, whereas in IF equalization $\mathbf{A}$ is only required to be full-rank. Second, the use of the dithers in IF equalization results in statistical independence between v_k and $z_{\text{eff},k}$ at each of the $2M$ sub-channels. This allows for an exact rigorous analysis of the error probability, which is seemingly difficult under the LR framework.

D. Bounding the Effective SNR for an optimal choice of $\mathbf{A}$

In this subsection, parts of the derivation from [12, Proof of Theorem 4] are followed in order to obtain a lower bound on $\mathsf{SNR}_{\text{eff}}$.

The target integer-valued matrix $\mathbf{A}$ should be chosen such as to maximize $\mathsf{SNR}_{\text{eff}}$. This criterion translates to choosing

$$\begin{aligned} \mathbf{A}^{\text{opt}} &= \underset{\mathbf{A}\in\mathbb{Z}^{2M\times 2M},\det(\mathbf{A})\neq 0}{\operatorname{argmax}} \mathsf{SNR}_{\text{eff}} \\ &= \underset{\mathbf{A}\in\mathbb{Z}^{2M\times 2M},\det(\mathbf{A})\neq 0}{\operatorname{argmin}} \max_{k=1\ldots,2M} \mathbf{a}_k^T\left(\mathbf{I}+\mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}\mathbf{a}_k. \end{aligned}$$

Consider the Cholesky decomposition of the matrix $\left(\mathbf{I}+\mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1}$ (such a decomposition exists since the matrix is symmetric and positive definite)

$$\left(\mathbf{I}+\mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)^{-1} = \mathbf{L}\mathbf{L}^T, \tag{11}$$

where $\mathbf{L}$ is a lower triangular matrix with strictly positive diagonal entries. With this notation the optimization criterion becomes

$$\mathbf{A}^{\text{opt}} = \underset{\mathbf{A}\in\mathbb{Z}^{2M\times 2M},\det(\mathbf{A})\neq 0}{\operatorname{argmin}} \max_{k=1\ldots,2M} \|\mathbf{L}^T\mathbf{a}_k\|^2.$$

Denote by $\Lambda(\mathbf{L}^T)$ the $2M$ dimensional lattice spanned by the matrix $\mathbf{L}^T$, i.e.,

$$\Lambda(\mathbf{L}^T) \triangleq \left\{\mathbf{L}^T\mathbf{a} \;:\; \mathbf{a}\in\mathbb{Z}^{2M}\right\}.$$

It follows that $\mathbf{A}^{\text{opt}}$ should consist of the set of $2M$ linearly independent integer-valued vectors that result in the shortest set of linearly independent lattice vectors in $\Lambda(\mathbf{L}^T)$.

Definition 1 (Successive minima): Let $\Lambda(\mathbf{L}^T)$ be a lattice spanned by the full-rank matrix $\mathbf{L}^T \in \mathbb{R}^{2M\times 2M}$. For $k = 1, \ldots, 2M$, we define the kth successive minimum as

$$\lambda_k(\mathbf{L}^T) \triangleq \inf\left\{r \;:\; \dim\left(\operatorname{span}\left(\Lambda(\mathbf{L}^T)\bigcap\mathcal{B}(\mathbf{0},r)\right)\right) \geq k\right\}$$

where $\mathcal{B}(\mathbf{0},r) = \left\{\mathbf{x}\in\mathbb{R}^{2M} \;:\; \|\mathbf{x}\| \leq r\right\}$ is the closed ball of radius r around $\mathbf{0}$. In words, the kth successive minimum of a lattice is the minimal radius of a ball centered around $\mathbf{0}$ that contains k linearly independent lattice points.

With the above definition of successive minima, the effective signal-to-noise ratio, when the optimal integer-valued matrix $\mathbf{A}^{\text{opt}}$ is used, can be written as

$$\mathsf{SNR}_{\text{eff}} = \frac{1}{\lambda^2_{2M}(\mathbf{L}^T)}. \tag{12}$$

Bounding the value of the $2M$th successive minima of a lattice is seemingly difficult. Fortunately, the theory of lattices connects the $2M$th successive minima of a lattice to the first successive minima of its dual lattice.

Definition 2 (Dual lattice): For a lattice $\Lambda(\mathbf{G})$ with a generating full-rank matrix $\mathbf{G} \in \mathbb{R}^{2M\times 2M}$ the dual lattice is defined by

$$\begin{aligned} \Lambda^*(\mathbf{G}) &\triangleq \Lambda\left((\mathbf{G}^T)^{-1}\right) \\ &= \left\{(\mathbf{G}^T)^{-1}\mathbf{a} \;:\; \mathbf{a}\in\mathbb{Z}^{2M}\right\}. \end{aligned}$$

Lemma 2 ([25, Proposition 3.3]): Let $\Lambda(\mathbf{G})$ be a lattice with a full-rank generating matrix $\mathbf{G} \in \mathbb{R}^{2M\times 2M}$ and let $\Lambda^*(\mathbf{G}) = \Lambda\left((\mathbf{G}^T)^{-1}\right)$ be its dual lattice. The successive minima of $\Lambda(\mathbf{G})$ and $\Lambda^*(\mathbf{G})$ satisfy the following inequality

$$\lambda^2_\ell(\mathbf{G})\,\lambda^2_1\left((\mathbf{G}^T)^{-1}\right) < (2M)^3, \quad \forall \ell = 1,2,\ldots,2M.$$

Proof: See [25] ∎

The dual lattice of $\Lambda(\mathbf{L}^T)$ is $\Lambda(\mathbf{L}^{-1})$. Applying Lemma 2 to (12) gives

$$\begin{aligned}\mathsf{SNR}_{\text{eff}} &= \frac{1}{\lambda_{2M}^2(\mathbf{L}^T)} \\ &> \frac{1}{(2M)^3}\lambda_1^2(\mathbf{L}^{-1}) \\ &= \frac{1}{8M^3}\min_{\mathbf{a}\in\mathbb{Z}^{2M}\setminus\mathbf{0}} \|\mathbf{L}^{-1}\mathbf{a}\|^2 \\ &= \frac{1}{8M^3}\min_{\mathbf{a}\in\mathbb{Z}^{2M}\setminus\mathbf{0}} \mathbf{a}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{a} \\ &= \frac{1}{8M^3}\min_{\mathbf{a}\in\mathbb{Z}^{2M}\setminus\mathbf{0}} \mathbf{a}^T\left(\mathbf{I}+\mathsf{SNR}\tilde{\mathbf{H}}^T\tilde{\mathbf{H}}\right)\mathbf{a}. \end{aligned} \tag{13}$$

where (13) follows from (11).

The performance of IF equalization over Rayleigh fading channels was studied in [12]. Using bounding techniques similar to those used in this subsection, it was shown that the IF equalizer achieves the optimal receive DMT (corresponding to transmission of independent streams from each antenna) when $N \geq M$. Nevertheless, there are clearly instances of MIMO channels for which the lower bound (13) on $\mathsf{SNR}_{\text{eff}}$ does not increase with the WI mutual information. For example, consider a channel $\mathbf{H}$ where one of the NM entries equals h whereas all other gains are zero. For such a channel $C_{\text{WI}} = \log(1+|h|^2\mathsf{SNR})$, yet $\mathsf{SNR}_{\text{eff}} = 1$ (and the bound (13) only gives $\mathsf{SNR}_{\text{eff}} > 1/(8M^3)$). Thus, it is evident that IF equalization alone can perform arbitrarily far from C_{WI}.

As shown in the sequel, this problem can be overcome by transmitting linear combinations of multiple streams from each antenna. More precisely, instead of transmitting $2M$ linearly coded streams, one from the in-phase component and one from the quadrature component of each antenna, over n channel uses, $2MT$ linearly coded streams are precoded by a unitary matrix and transmitted over nT channel uses.

E. Linearly Precoded Integer-Forcing Equalization

Very recently Domanovitz *et al.* [26] proposed to combine IF equalization with linear precoding. The idea is to transform the $N \times M$ complex MIMO channel (1) into an aggregate $NT \times MT$ complex MIMO channel and then apply IF equalization to the aggregate channel. The transformation is done using a unitary precoding matrix $\mathbf{P} \in \mathbb{C}^{MT\times MT}$. Specifically, let $\bar{\mathbf{x}} \in \mathbb{C}^{MT\times 1}$ be the input vector to the aggregate channel. This vector is multiplied by $\mathbf{P}$ to form the vector $\mathbf{x} = \mathbf{P}\bar{\mathbf{x}} \in \mathbb{C}^{MT\times 1}$ which is transmitted over the channel (1) during T consecutive channel uses. Let

$$\mathcal{H} = \mathbf{I}_T \otimes \mathbf{H} = \begin{bmatrix} \mathbf{H} & 0 & \cdots & 0 \\ 0 & \mathbf{H} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{H} \end{bmatrix}, \tag{14}$$

where $\otimes$ denotes the Kronecker product. The output of the aggregate channel is obtained by stacking T consecutive outputs of the channel (1) one below the other and is given by

$$\begin{aligned}\bar{\mathbf{y}} &= \mathcal{H}\mathbf{P}\bar{\mathbf{x}} + \bar{\mathbf{z}} \\ &= \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{z}}, \end{aligned}\tag{15}$$

where $\bar{\mathbf{H}} \triangleq \mathcal{H}\mathbf{P} = (\mathbf{I}_T \otimes \mathbf{H})\mathbf{P} \in \mathbb{C}^{NT\times MT}$ is the aggregate channel matrix, and $\bar{\mathbf{z}} \in \mathbb{C}^{NT\times 1}$ is a vector of i.i.d. circularly symmetric complex normal noise. See Figure 3.

A remaining major challenge is how to choose the precoding matrix $\mathbf{P}$ (recall that an open-loop scenario is considered, and hence the choice of $\mathbf{P}$ cannot depend on $\mathbf{H}$). As observed in Section II-C, the performance of the IF equalizer is dictated by $\mathsf{SNR}_{\text{eff}}$. Thus, in order to obtain achievable rates that are comparable to the WI mutual information, $\mathsf{SNR}_{\text{eff}}$ must scale appropriately with C_{WI}. The precoding matrix $\mathbf{P}$ should be chosen such as to guarantee this property for all channel matrices with the same WI mutual information.

It turns out that the well-developed theory of linear dispersion space-time codes provides excellent candidates for $\mathbf{P}$ that satisfy the aforementioned criterion. In [26] the performance of IF equalization with the golden code's [11] precoding matrix was numerically evaluated in a 2×2 MIMO Rayleigh fading environment. The scheme's outage probability was found to be relatively close to that achieved by white i.i.d. Gaussian codebooks. Here, we prove that, in fact, precoded IF equalization, where the precoding matrix generates a perfect linear dispersion space-time code, achieves rates within a *constant gap* from the WI mutual information of *any* MIMO channel.

The next section provides some necessary background on space-time codes, needed for the proof of our main result.

III. Linear Dispersion Space-Time Codes

An $M\times T$ space-time (ST) code $\mathcal{C}^{\text{ST}}$ for the channel (1) with rate R is a set of $|\mathcal{C}^{\text{ST}}| = 2^{RT}$ complex matrices of dimensions $M \times T$. The codebook $\mathcal{C}^{\text{ST}}$ has to satisfy the average power constraint[5]

$$\frac{1}{2^{RT}}\sum_{\mathbf{X}_k\in\mathcal{C}^{\text{ST}}} \|\mathbf{X}_k\|_F^2 \leq MT\cdot\mathsf{SNR}.$$

When the ST code $\mathcal{C}^{\text{ST}}$ is used, a code matrix $\mathbf{X} \in \mathcal{C}^{\text{ST}}$ is transmitted column by column over T consecutive channel uses, such that the T channel outputs can be expressed as

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z},$$

where each column of the matrices $\mathbf{Y}, \mathbf{Z} \in \mathbb{C}^{N\times T}$ represents the channel output and additive noise, respectively, at one of the T channel uses.

An ST code $\mathcal{C}^{\text{ST}}$ is said to be a *linear dispersion* ST code [27] over the constellation $\mathcal{S}$ if every code matrix $\mathbf{X} \in \mathcal{C}^{\text{ST}}$ can be uniquely decomposed as

$$\mathbf{X} = \sum_{k=1}^{K} s_k\mathbf{F}_k,\ \ s_k \in \mathcal{S},$$

where $\mathcal{S}$ is some constellation and the matrices $\mathbf{F}_k \in \mathbb{C}^{M\times T}$ are fixed and independent of the constellation symbols s_k. Denoting by $\text{vec}(\mathbf{X})$ the vector obtained by stacking the columns of $\mathbf{X}$ one below the other, and letting $\mathbf{s} = [s_1\ \cdots\ s_K]^T$ gives

$$\text{vec}(\mathbf{X}) = \mathbf{P}\mathbf{s},$$

[5]The Forbenius norm of a matrix $\mathbf{X}$ is denoted by $\|\mathbf{X}\|_F^2$.

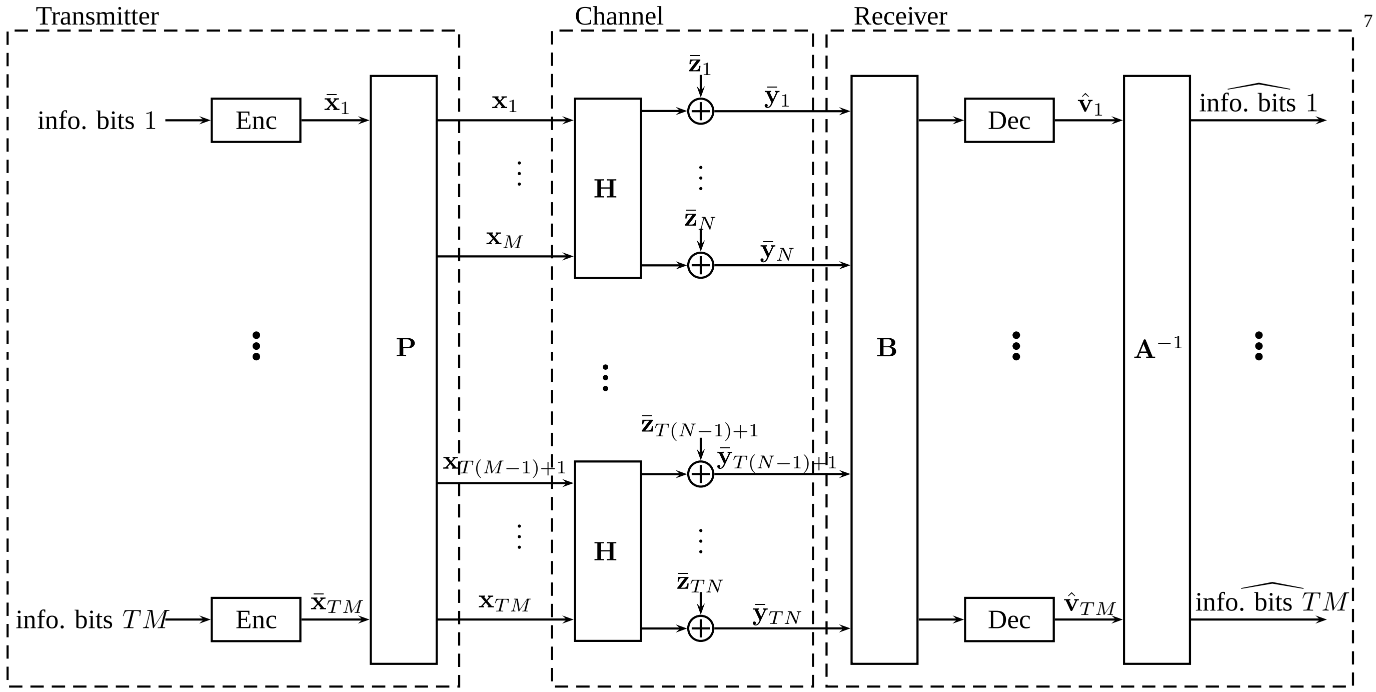

Fig. 3. A schematic overview of linearly precoded integer-forcing. For simplicity, the dithers are not depicted in the figure, and a real-valued channel is assumed. At the transmitter, the information bits are split to TM streams, each of which is encoded by a similar linear code. Then, a $TM \times TM$ precoding matrix "mixes" the TM codewords into TM linear combinations. The channel $\mathbf{H}$ is used T times, where in each channel use each of the antennas transmits one of the precoded linear combinations. The receiver treats T consecutive channel outputs as the output of an aggregate $NT \times MT$ channel with transfer matrix $\bar{\mathbf{H}} = (\mathbf{I}_T \otimes \mathbf{H})\mathbf{P}$, and applies integer-forcing equalization to the aggregate channel.

where

$$\mathbf{P} = [\text{vec}(\mathbf{F}_1) \;\; \text{vec}(\mathbf{F}_2) \;\; \cdots \;\; \text{vec}(\mathbf{F}_K)]$$

is the code's $MT \times K$ *generating matrix*. A linear dispersion ST code is *full-rate* if $K = MT$. In the sequel, linear dispersion codes over an $L^2 -$QAM constellation will play a key role. Here, an $L^2 -$ QAM constellation is defined as

$$\{-L, -L+1, \ldots, L-1, L\} + i\,\{-L, -L+1, \ldots, L-1, L\}$$

for some positive integer L. The linear dispersion ST code obtained by using the infinite constellation $\mathcal{S}_\infty = \mathbb{Z} + i\mathbb{Z}$ is referred to as $\mathcal{C}_\infty^{\text{ST}}$, and, after vectorization, is in fact a complex lattice with generating matrix $\mathbf{P}$. Since the L^2-QAM constellation is a subset of $\mathbb{Z}+i\mathbb{Z}$ it follows that for any finite L the $L^2 -$ QAM based code $\mathcal{C}^{\text{ST}}$ is a subset of $\mathcal{C}_\infty^{\text{ST}}$.

An important class of linear dispersion ST codes is that of *perfect codes* [8], [9] which is defined next.

Definition 3: An $M \times M$ linear dispersion ST code over a QAM constellation is called *perfect* if

1) It is full-rate;
2) It satisfies the nonvanishing determinant criterion

$$\delta_{\min}(\mathcal{C}_\infty^{\text{ST}}) = \min_{\mathbf{0}\neq\mathbf{X}\in\mathcal{C}_\infty^{\text{ST}}} |\det(\mathbf{X})|^2 > 0;$$

3) The code's generating matrix is unitary, i.e., $\mathbf{P}^\dagger\mathbf{P} = \mathbf{I}$.

Note that this definition is slightly different than the one used in [8], [9], where instead of condition 3 it is required that the energy of the codeword corresponding to the information symbols s will have the same energy as $\|\mathbf{s}\|^2$, and that all the coded symbols in all T time slots will have the same average energy.

In [8], perfect linear dispersion ST codes were found for $M = 2, 3, 4$ and 6, whereas in [9] perfect linear dispersion ST codes were obtained for any positive integer M. The constructions in [8], [9] are based on cyclic division algebras, and results in a unitary generating matrix $\mathbf{P}$. Thus, for any positive integer M, there exist codes that are perfect according to Definition 3.

The approximate universality of an ST code over the MIMO channel was studied in [6]. This property refers to an ST code being optimal in terms of DMT regardless of the fading statistics of $\mathbf{H}$. A sufficient and necessary condition for an ST code to be approximately universal was derived in [6, Theorem 3.1]. This condition is closely related to the nonvanishing determinant criterion. The next Theorem is a simple extension of [6, Theorem 3.1]. The notation $[x]^+ \triangleq \max(x, 0)$ is used.

Theorem 1: Let $\mathcal{C}^{\text{ST}}$ be an $M \times M$ perfect linear dispersion ST code over an $L^2 -$ QAM constellation with $\delta_{\min}(\mathcal{C}_\infty^{\text{ST}}) = \min_{\mathbf{0}\neq\mathbf{X}\in\mathcal{C}_\infty^{\text{ST}}} |\det(\mathbf{X})|^2 > 0$. Then, for all channel matrices $\mathbf{H}$ with corresponding WI mutual information $C_{\text{WI}} = \log\det(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})$ and all $\mathbf{0} \neq \mathbf{X} \in \mathcal{C}^{\text{ST}}$

$$\mathsf{SNR}\|\mathbf{H}\mathbf{X}\|_F^2 \geq \left[\delta_{\min}(\mathcal{C}_\infty^{\text{ST}})^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}} - 2M^2L^2\right]^+.$$

The proof closely follows that of [6, Theorem 3.1].

Proof: Consider some arbitrary $\mathbf{0} \neq \mathbf{X} \in \mathcal{C}^{\text{ST}}$ and let

$$\mathbf{H} = \mathbf{U}_1\mathbf{\Psi}\mathbf{V}_1^\dagger \text{ and } \mathbf{X} = \mathbf{U}_2\mathbf{\Lambda}\mathbf{V}_2^\dagger$$

be the singular value decompositions (SVD) of $\mathbf{H}$ and $\mathbf{X}$, respectively. With this notation

$$\mathsf{SNR}\|\mathbf{H}\mathbf{X}\|_F^2 = \mathsf{SNR}\|\mathbf{\Psi}\mathbf{V}_1^\dagger\mathbf{U}_2\mathbf{\Lambda}\|_F^2. \tag{16}$$

Suppose the (absolute) singular values are ordered by increasing value in $\mathbf{\Lambda}$ and by decreasing value in $\mathbf{\Psi}$:

$$\begin{aligned}\mathbf{\Lambda} &= \mathrm{diag}\{\lambda_1, \ldots, \lambda_M\},\\ \mathbf{\Psi} &= \mathrm{diag}\{\psi_1, \ldots, \psi_{m_n}, 0, \cdots, 0\},\end{aligned}$$

where $m_n \triangleq \min\{M, N\}$. In order to establish the desired result one has to find the channel $\mathbf{H}$ with corresponding WI mutual information C_{WI} that minimizes (16). The rotation matrix $\mathbf{V}_1$ that minimizes (16) is $\mathbf{V}_1 = \mathbf{U}_2$ which aligns the weaker singular values of the channel matrix with the stronger singular values of the code matrix [28]. Thus, the problem of finding the worst channel matrix $\mathbf{H}$ w.r.t. the codeword $\mathbf{X}$ reduces to the optimization problem

$$\begin{aligned}&\min_{\psi_1,\ldots,\psi_{m_n}} \mathsf{SNR} \sum_{m=1}^{m_n} |\psi_m|^2 |\lambda_m|^2\\ &\text{subject to} \sum_{m=1}^{m_n} \log(1 + |\psi_m|^2 \mathsf{SNR}) = C_{\mathrm{WI}}.\end{aligned} \tag{17}$$

A lower bound on the solution of the minimization problem (17) can be obtained by replacing m_n with $M \geq m_n$, which increases the optimization space and results in

$$\begin{aligned}&\min_{\psi_1,\ldots,\psi_{M}} \mathsf{SNR} \sum_{m=1}^{M} |\psi_m|^2 |\lambda_m|^2\\ &\text{subject to} \sum_{m=1}^{M} \log(1 + |\psi_m|^2 \mathsf{SNR}) = C_{\mathrm{WI}}.\end{aligned} \tag{18}$$

The solution to (18) is given by standard water-filling [6]

$$\mathsf{SNR}\|\mathbf{H}\mathbf{X}\|_F^2 \geq \sum_{m=1}^{M} \left[\frac{1}{\lambda} - |\lambda_m|^2\right]^+, \tag{19}$$

where λ satisfies

$$\sum_{m=1}^{M} \left[\log\left(\frac{1}{\lambda|\lambda_m|^2}\right)\right]^+ = C_{\mathrm{WI}}. \tag{20}$$

Recall that $\mathbf{X}$ is a codeword from a perfect linear dispersion ST code over an $L^2 -$ QAM constellation. Let $\mathbf{P}$ be the generating matrix of the code $\mathcal{C}^{\mathrm{ST}}$. Thus, $\mathrm{vec}(\mathbf{X}) = \mathbf{P}\mathbf{s}$ for some vector $\mathbf{s}$ whose M^2 components all belong to the $L^2 -$ QAM constellation. This implies that

$$\begin{aligned}\sum_{m=1}^{M} |\lambda_m|^2 &= \|\mathbf{X}\|_F^2\\ &= \|\mathrm{vec}(\mathbf{X})\|^2\\ &= \|\mathbf{P}\mathbf{s}\|^2\\ &= \|\mathbf{s}\|^2 &(21)\\ &\leq 2M^2L^2, &(22)\end{aligned}$$

where (21) follows from the fact that $\mathbf{P}$ is unitary. In particular, (22) implies that

$$|\lambda_m|^2 \leq 2M^2L^2$$

for all $m = 1, \ldots, M$. Without loss of generality we may assume that $2M^2L^2 \leq \delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}}$ as otherwise the theorem is trivial. Since by definition

$$|\lambda_1 \cdots \lambda_M|^2 = |\det(\mathbf{X})|^2 \geq \delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}}),$$

we have for all $m = 1, \ldots, M$

$$\begin{aligned}|\lambda_m|^2 &\leq 2M^2L^2\\ &\leq \delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}}\\ &\leq |\lambda_1 \cdots \lambda_M|^{\frac{2}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}}.\end{aligned} \tag{23}$$

Thus, the solution for (20) is

$$\frac{1}{\lambda} = |\lambda_1 \cdots \lambda_M|^{\frac{2}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}}. \tag{24}$$

Substituting (24) into (19) gives

$$\begin{aligned}\mathsf{SNR}\|\mathbf{H}\mathbf{X}\|_F^2 &\geq \left[M|\lambda_1 \cdots \lambda_M|^{\frac{2}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}} - \sum_{m=1}^{M} |\lambda_m|^2\right]^+\\ &\geq \left[M\delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}} - 2M^2L^2\right]^+\\ &\geq \left[\delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}} - 2M^2L^2\right]^+\end{aligned}$$

as desired. ■

Let $\mathcal{H} = \mathbf{I}_M \otimes \mathbf{H}$, as in (14). The next simple corollary of Theorem 1 will be used in Section IV to prove the main result of this paper.

Corollary 1: Let $\mathbf{P}$ be the $M^2 \times M^2$ generating matrix of a perfect linear dispersion code $\mathcal{C}^{\mathrm{ST}}$ over a QAM constellation with $\delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}}) = \min_{\mathbf{0} \neq \mathbf{X} \in \mathcal{C}_\infty^{\mathrm{ST}}} |\det(\mathbf{X})|^2 > 0$. Then, for all channel matrices $\mathbf{H}$ with corresponding white input capacity $C_{\mathrm{WI}} = \log\det(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})$ and all complex vectors $\mathbf{a} \in \mathbb{Z}^{M^2} + i\mathbb{Z}^{M^2}$ with all real and imaginary entries not greater than L

$$\mathsf{SNR}\|\mathcal{H}\mathbf{P}\mathbf{a}\|^2 \geq \left[\delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}} - 2M^2L^2\right]^+.$$

Proof: Assume the constellation over which the code $\mathcal{C}^{\mathrm{ST}}$ is defined is an $L^2 -$ QAM one. Then, for any $\mathbf{a} \in \mathbb{Z}^{M^2} + i\mathbb{Z}^{M^2}$ with all real and imaginary entries not greater than L there exist a code matrix $\mathbf{X} \in \mathcal{C}^{\mathrm{ST}}$ such that

$$\mathrm{vec}(\mathbf{X}) = \mathbf{P}\mathbf{a}.$$

Now,

$$\begin{aligned}\mathsf{SNR}\|\mathcal{H}\mathbf{P}\mathbf{a}\|^2 &= \mathsf{SNR}\|\mathcal{H}\,\mathrm{vec}(\mathbf{X})\|^2\\ &= \mathsf{SNR}\|\mathbf{H}\mathbf{X}\|_F^2\\ &\geq \left[\delta_{\mathrm{min}}(\mathcal{C}_\infty^{\mathrm{ST}})^{\frac{1}{M}} 2^{\frac{C_{\mathrm{WI}}}{M}} - 2M^2L^2\right]^+,\end{aligned}$$

where the last inequality follows from Theorem 1. ■

IV. Main Result

The next theorem lower bounds the effective signal-to-noise ratio of linearly precoded IF equalization, where the precoding matrix generates a perfect linear dispersion ST code. The obtained bound depends on the channel matrix $\mathbf{H}$ *only* through its corresponding WI mutual information.

Theorem 2: Consider the aggregate MIMO channel

$$\bar{\mathbf{y}} = \mathcal{H}\mathbf{P}\bar{\mathbf{x}} + \bar{\mathbf{z}}$$

where $\mathcal{H} = \mathbf{I}_M \otimes \mathbf{H} \in \mathbb{C}^{NM \times M^2}$, and $\mathbf{P} \in \mathbb{C}^{M^2 \times M^2}$ is a generating matrix of a perfect $M \times M$ QAM based ST code $\mathcal{C}^{\text{ST}}$ with $\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty) = \min_{\mathbf{0}\neq\mathbf{X}\in\mathcal{C}^{\text{ST}}_\infty} |\det(\mathbf{X})|^2 > 0$. Then, when IF equalization is applied to the aggregate channel

$$\mathsf{SNR}_{\text{eff}} > \frac{1}{16M^8}\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}},$$

for all channel matrices $\mathbf{H}$ with corresponding WI mutual information $C_{\text{WI}} = \log\det(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})$.

Proof: Let $\bar{\mathbf{H}} = \mathcal{H}\mathbf{P}$ and consider its real-valued representation

$$\tilde{\bar{\mathbf{H}}} \triangleq \begin{bmatrix} \bar{\mathbf{H}}_{\Re} & -\bar{\mathbf{H}}_{\Im} \\ \bar{\mathbf{H}}_{\Im} & \bar{\mathbf{H}}_{\Re} \end{bmatrix} \in \mathbb{R}^{2NM \times 2M^2}.$$

It follows from (13) that the effective-signal-to-noise ratio when IF equalization is applied to the aggregate channel is upper bounded by

$$\mathsf{SNR}_{\text{eff}} > \frac{1}{8M^6} \min_{\mathbf{a}\in\mathbb{Z}^{2M^2}\setminus\mathbf{0}} \mathbf{a}^T \left(\mathbf{I} + \mathsf{SNR}\tilde{\bar{\mathbf{H}}}^T\tilde{\bar{\mathbf{H}}}\right)\mathbf{a}. \tag{25}$$

Since the matrix $\left(\mathbf{I} + \mathsf{SNR}\tilde{\bar{\mathbf{H}}}^T\tilde{\bar{\mathbf{H}}}\right) \in \mathbb{R}^{2M^2 \times 2M^2}$ is the real-valued representation of the complex matrix $\left(\mathbf{I} + \mathsf{SNR}\bar{\mathbf{H}}^T\bar{\mathbf{H}}\right) \in \mathbb{C}^{M^2 \times M^2}$, (25) can be written in complex form as

$$\begin{aligned} \mathsf{SNR}_{\text{eff}} &> \frac{1}{8M^6} \min_{\mathbf{a}\in\mathbb{Z}^{M^2}+i\mathbb{Z}^{M^2}\setminus\mathbf{0}} \mathbf{a}^T\left(\mathbf{I} + \mathsf{SNR}\bar{\mathbf{H}}^T\bar{\mathbf{H}}\right)\mathbf{a} \\ &= \frac{1}{8M^6} \min_{\mathbf{a}\in\mathbb{Z}^{M^2}+i\mathbb{Z}^{M^2}\setminus\mathbf{0}} \left(\|\mathbf{a}\|^2 + \mathsf{SNR}\|\bar{\mathbf{H}}\mathbf{a}\|^2\right) \\ &= \frac{1}{8M^6} \min_{\mathbf{a}\in\mathbb{Z}^{M^2}+i\mathbb{Z}^{M^2}\setminus\mathbf{0}} \left(\|\mathbf{a}\|^2 + \mathsf{SNR}\|\mathcal{H}\mathbf{P}\mathbf{a}\|^2\right). \end{aligned} \tag{26}$$

For a vector $\mathbf{a}' \in \mathbb{Z}^{M^2} + i\mathbb{Z}^{M^2} \setminus \mathbf{0}$ assume that the maximum absolute value of all its real and complex entries is L, i.e.,

$$L = \max\left[|\mathbf{a}'_{\Re}| \ |\mathbf{a}'_{\Im}|\right].$$

The norm $\|\mathbf{a}'\|^2$ is trivially bounded from below by L^2 and by Corollary 1

$$\mathsf{SNR}\|\mathcal{H}\mathbf{P}\mathbf{a}'\|^2 \geq \left[\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}} - 2M^2L^2\right]^+.$$

Therefore, (26) can be further bounded by

$$\begin{aligned} \mathsf{SNR}_{\text{eff}} &> \frac{1}{8M^6} \min_{L=1,2,\dots} \left(L^2 + \left[\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}} - 2M^2L^2\right]^+\right) \\ &\geq \frac{1}{8M^6} \min_{L=1,2,\dots} \left(L^2 + \left[\frac{\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}}}{2M^2} - L^2\right]^+\right) \\ &\geq \frac{1}{16M^8}\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}} \end{aligned}$$

as desired. ■

The next theorem is the main result of this paper.

Theorem 3: For all channel matrices $\mathbf{H}$ with M transmit antennas and an arbitrary number of receive antennas, the achievable rate of linearly precoded IF equalization is a constant gap from the closed-loop capacity of the channel, provided that the precoding matrix $\mathbf{P}$ generates an $M \times M$ perfect linear dispersion space-time code.

Proof: Let $C_{\text{WI}} = \log\det(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})$ be the WI mutual information corresponding to $\mathbf{H}$, and let $\mathbf{P}$ be a generating matrix of a perfect $M \times M$ QAM based code $\mathcal{C}^{\text{ST}}$ with $\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty) = \min_{\mathbf{0}\neq\mathbf{X}\in\mathcal{C}^{\text{ST}}_\infty} |\det(\mathbf{X})|^2 > 0$. Such a matrix $\mathbf{P}$ exists for any positive integer M [9]. The matrix $\mathbf{P}$ is used as a precoding matrix that transforms the original $N \times M$ MIMO channel (1) to the aggregate $NM \times M^2$ MIMO channel

$$\begin{aligned} \bar{\mathbf{y}} &= \mathcal{H}\mathbf{P}\bar{\mathbf{x}} + \bar{\mathbf{z}} \\ &= \bar{\mathbf{H}}\bar{\mathbf{x}} + \bar{\mathbf{z}}, \end{aligned} \tag{27}$$

as described in Section II-E, and then IF equalization is applied to the aggregate channel. Assuming a "good" nested lattice codebook is used to encode all $2M^2$ streams transmitted over the aggregate channel, by (8), IF equalization can achieve any rate satisfying

$$R_{\text{IF,aggregate}} < M^2\log(\mathsf{SNR}_{\text{eff}}).$$

Using Theorem 2, it follows that any rate satisfying

$$\begin{aligned} R_{\text{IF,aggregate}} &< M^2\log\left(\frac{1}{16M^8}\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)^{\frac{1}{M}} 2^{\frac{C_{\text{WI}}}{M}}\right) \\ &= MC_{\text{WI}} - M\log\frac{1}{\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)} - M^2\log(16M^8) \end{aligned} \tag{28}$$

is achievable over the aggregate channel.

Since each channel use of the aggregate channel (27) corresponds to M channel uses of the original channel (1), the communication rate should be normalized by a factor of $1/M$. Thus, linearly precoded IF equalization can achieve any rate satisfying

$$\begin{aligned} R_{\text{P-IF}} &< C_{\text{WI}} - \log\frac{1}{\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)} - M\log(16M^8) \\ &= C_{\text{WI}} - \Gamma\left(\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty), M\right), \end{aligned}$$

where

$$\Gamma\left(\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty), M\right) \triangleq \log\frac{1}{\delta_{\min}(\mathcal{C}^{\text{ST}}_\infty)} + M\log(16M^8) \tag{29}$$

is a constant independent of the channel $\mathbf{H}$ and SNR. Since the gap between C_{WI} and the closed-loop capacity is bounded by a constant number of bits, the theorem is established. ■

Note that although the proof of Theorem 3 assumed that a "good" nested lattice code was used, a similar result holds when a q-ary linear code without shaping is used. This follows from the fact that the latter code loses only the shaping loss of $\log(2\pi e/12)$ bits per antenna w.r.t. the former. Moreover, Theorem 2 can also be used to obtain an upper bound on the

performance of linearly precoded IF equalization with uncoded transmission.

Proposition 1: For all channel matrices $\mathbf{H}$ with corresponding WI mutual information $C_{\text{WI}} = \log\det(\mathbf{I} + \mathsf{SNR}\mathbf{H}^\dagger\mathbf{H})$, M transmit antennas and an arbitrary number of receive antennas, the error probability of linearly precoded IF equalization with uncoded transmission is bounded by

$$P_{e,\text{P-IF-uncoded}} \leq 4M^2 \exp\left\{-\frac{3}{2}2^{\frac{1}{M}\left(C_{\text{WI}} - R_{\text{P-IF}} - \Gamma\left(\delta_{\min}(\mathcal{C}_\infty^{\text{ST}}), M\right)\right)}\right\}$$

provided that the precoding matrix $\mathbf{P}$ generates an $M \times M$ perfect linear dispersion space-time code $\mathcal{C}^{\text{ST}}$ with $\delta_{\min}(\mathcal{C}_\infty^{\text{ST}}) = \min_{\mathbf{0}\neq\mathbf{X}\in\mathcal{C}_\infty^{\text{ST}}} |\det(\mathbf{X})|^2 > 0$.

Proof: Using (10), the error probability of uncoded IF equalization over the aggregate channel (27) is bounded by

$$P_{e,\text{P-IF-uncoded}} \leq 4M^2 \exp\left\{-\frac{3}{2}2^{\frac{1}{M^2}\left(M^2\log(\mathsf{SNR}_{\text{eff}}) - MR_{\text{P-IF}}\right)}\right\},$$

where we have used the fact that the transmission rate over the aggregate channel is M times larger than the actual communication rate $R_{\text{P-IF}}$. Now, replacing $\mathsf{SNR}_{\text{eff}}$ with its bound from Theorem 2 establishes the proposition. ■

V. Discussion and Summary

The additive Gaussian noise MIMO channel in an open-loop scenario, where the receiver has complete channel state information whereas the transmitter has no channel state information was considered in this paper. It was shown that using linear precoding at the transmitter in conjunction with integer-forcing equalization at the receiver suffices to approach the closed-loop capacity of this channel to within a constant gap, regardless of the channel matrix $\mathbf{H}$. To the best of our knowledge, this is the first practical scheme that guarantees only an additive loss w.r.t. capacity. Such a performance guarantee is much stronger than DMT optimality, which is at present the common benchmark for evaluating schemes. In particular, although our results are free from any statistical assumptions, they can be interpreted to obtain performance guarantees in a MIMO fading environment. Specifically, a scheme that achieves a constant gap from capacity is DMT optimal under any fading statistics, and achieves a constant gap from the outage capacity under any fading statistics.

IF equalization uses coded streams, and is therefore usually less suitable for fast fading environments. Nevertheless, we have also developed new upper bounds on an uncoded version of IF equalization, which is more adequate for fast fading. We note that while uncoded IF equalization is quite similar to lattice reduction aided decoding, to the best of our knowledge, the performance of the latter was never studied at such a fine scale.

Another appealing feature of the described scheme is that is is independent of the number of receive antennas, and the performance guarantees obtained in this paper do not depend on the number of receive antennas as well. Hence, the scheme is not sensitive to a degrees-of-freedom mismatch.

Acknowledgment

Helpful discussions with Yair Yona, Elad Domanovitz and Bobak Nazer are greatly acknowledged.

Appendix A
Proof of Lemma 1

The probability density function of z_{eff} is symmetric around zero and hence

$$\Pr(z_{\text{eff}} \geq \tau) = \Pr(z_{\text{eff}} \leq -\tau).$$

Applying the Chernoff bound gives (for $s > 0$)

$$\begin{aligned}\Pr(z_{\text{eff}} \geq \tau) &\leq e^{-s\tau}\mathbb{E}\left(e^{sz_{\text{eff}}}\right) \\ &= e^{-s\tau}\mathbb{E}\left(e^{s\left(\sum_{\ell=1}^{L}\alpha_\ell z_l + \sum_{k=1}^{K}\beta_k d_k\right)}\right) \\ &= e^{-s\tau}\prod_{\ell=1}^{L}\mathbb{E}\left(e^{s\alpha_\ell z_l}\right)\prod_{k=1}^{K}\mathbb{E}\left(e^{s\beta_k d_k}\right).\end{aligned}$$

Using the well-known expressions for the moment generating function of Gaussian and uniform random variables gives

$$\begin{aligned}\mathbb{E}\left(e^{s\alpha_\ell z_l}\right) &= e^{\frac{1}{2}s^2\alpha_\ell^2\sigma_z^2}, \\ \mathbb{E}\left(e^{s\beta_k d_k}\right) &= \frac{\sinh(s\beta_k\rho/2)}{s\beta_k\rho/2} \leq e^{\frac{1}{2}\frac{s^2\beta_k^2\rho^2}{12}},\end{aligned}$$

where the last inequality follows from $\sinh(x)/x \leq \exp\{x^2/6\}$ (which can be obtained by simple Taylor expansion) [15]. It follows that

$$\begin{aligned}\Pr(z_{\text{eff}} \geq \tau) &\leq e^{-s\tau}e^{\frac{s^2}{2}\left(\sum_{\ell=1}^{L}\alpha_\ell^2\sigma_z^2 + \sum_{k=1}^{K}\beta_k^2\frac{\rho^2}{12}\right)} \\ &= e^{-s\tau + \frac{1}{2}s^2\sigma_{\text{eff}}^2}.\end{aligned} \tag{30}$$

Setting $s = \tau/\sigma_{\text{eff}}^2$ gives the desired result.

References

[1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Summer 1996.

[2] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, March 1998.

[3] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November - December 1999.

[4] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on Information Theory*, vol. 49, no. 5, pp. 1073–1096, may 2003.

[5] T. Philosof and R. Zamir, "The capacity loss of uncorrelated equi-power Gaussian input over MIMO channel," in *Proceedings of the 23rd IEEE Convention of Electrical and Electronics Engineerings in Israel (IEEEI 2004)*, Eilat, Israel, Sept., pp. 227–229.

[6] S. Tavildar and P. Viswanath, "Approximately universal codes over slow-fading channels," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3233–3258, july 2006.

[7] P. Elia, K. Kumar, S. Pawar, P. Kumar, and H.-F. Lu, "Explicit space-time codes achieving the diversity-multiplexing gain tradeoff," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3869–3884, Sept. 2006.

[8] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space-time block codes," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3885–3902, Sept. 2006.

[9] P. Elia, B. Sethuraman, and P. Vijay Kumar, "Perfect space-time codes for any number of antennas," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 3853–3868, Nov. 2007.

[10] H. Yao and G. W. Wornell, "Achieving the full MIMO diversity-multiplexing frontier with rotation-based space-time codes," in *Proceedings of the Allerton Conference on Communications, Control, and Computing*, 2003.

[11] J.-C. Belfiore, G. Rekaya, and E. Viterbo, "The golden code: a 2×2 full-rate space-time code with nonvanishing determinants," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1432 – 1436, Apr. 2005.

[12] J. Zhan, B. Nazer, U. Erez, and M. Gastpar, "Integer-forcing linear receivers," *IEEE Transactions on Information Theory*, Submitted January 2012, available online: http://arxiv.org/abs/1003.5966.

[13] J. Jalden and P. Elia, "DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models," *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4765–4780, Oct. 2010.

[14] U. Erez and R. Zamir, "Achieving $\frac{1}{2}\log\left(1+\mathrm{SNR}\right)$ on the AWGN channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, Oct. 2004.

[15] C. Feng, D. Silva, and F. Kschischang, "An algebraic approach to physical-layer network coding," *IEEE Transactions on Information Theory*, Submitted July 2011, available online http://arxiv.org/abs/1108.1695.

[16] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*. New York: Springer-Verlag, 1988.

[17] H.-A. Loeliger, "Averaging bounds for lattices and linear codes," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1767–1773, Nov. 1997.

[18] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3417–3432, Oct. 2005.

[19] J. Zhan, B. Nazer, O. Ordentlich, U. Erez, and M. Gastpar, "Integer-forcing architectures for MIMO: Distributed implementation and SIC," in *Proceedings of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Nov. 2010, pp. 322–326.

[20] O. Ordentlich, J. Zhan, U. Erez, B. Nazer, and M. Gastpar, "Practical code design for compute-and-forward," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT 2011)*, St. Petersburg, Russia, Aug. 2011.

[21] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[22] O. Ordentlich and U. Erez, "A simple proof for the existence of "good" pairs of nested lattices," in *Proceedings of the 27th Convention of Electrical Electronics Engineers in Israel (IEEEI)*, Nov. 2012, pp. 1–12.

[23] H. Yao and G. W. Wornell, "Lattice-reduction-aided detectors for MIMO communication systems," in *Proceedings of IEEE Globecom 2002*, Taipei, Taiwan, Nov. 2002.

[24] C. Windpassinger and R. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction," in *Proceedings of the Information Theory Workshop*, Apr. 2003, pp. 345–348.

[25] J. Lagarias, H. Lenstra Jr., and C. Schnorr, "Korkin-Zolotarev bases and successive minima of a lattice and its reciprocal lattice," *Combinatorica*, vol. 10, no. 4, pp. 333–348, Dec. 1990.

[26] E. Domanovitz and U. Erez, "Combining space-time block modulation with integer forcing receivers," in *Proceedings of the 27th Convention of Electrical Electronics Engineers in Israel (IEEEI)*, Nov. 2012, pp. 1–4.

[27] B. Hassibi and B. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[28] C. Kose and R. Wesel, "Universal space-time trellis codes," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2717–2727, Oct. 2003.